

SEQUENCE LISTING

RECEIVED

SEP 11 2001

TECH CENTER 1600/2900

<110> FLANNERY, ANGELA VERONICA
FINNEGAN, MARIA CHRISTINA MARTINA

<120> ZGGBP1, NOVEL PEPTIDES RELATED TO BIPOLAR AFFECTIVE
DISORDERS TYPE 1, SEQUENCES AND USES THEREOF

<130> DJB/009901/0265461

<140> 09/463,844

<141> 2000-02-01

<150> PCT/GB98/022059

<151> 1998-07-28

<150> GB 9716162.4

<151> 1997-08-01

<160> 17

<170> PatentIn Ver. 2.1

<210> 1

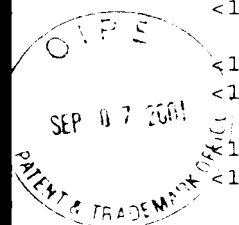
<211> 5154

<212> DNA

<213> Homo sapiens

<400> 1

caagcgcgca	attaaccctc	actaaaggga	acaccaacac	gtcgccagga	ctgcgccgtt	60
cgctgcgctc	ataggcgggc	atttcatcaa	gggtggcaag	gatcgccctg	tcgacgggtca	120
ggctgccttc	gacgcgggtt	ccctcctcgt	cctgttccag	ggtgagtggt	cgataccagg	180
tgtccaccgg	gaaggtacgg	cccgacacct	cgacaatcgg	cgcacgcgtc	aagtgccttg	240
aaaagcgctc	caggtcgatg	gtggccgagg	tgatgatgac	tttcaggctc	gggcgacgcg	300
gcaacagggg	cttgaggtag	ccgagcagga	agtcgatgtt	caggctgcgt	tcgtgggctt	360
cgtcgacgac	aggctcgcgt	tatggctccg	ctttctgcgg	ctctcctacc	ctggcatggt	420
gtgtgtgtgt	gcctgtgtgc	tacggagagt	cccgatttct	cagagtaaaa	gttgttcttg	480
aatgatctcg	ccaaaaagga	catctttgga	gccagtgatc	cgtatgtgaa	actttcattg	540
tacgtagcgg	atgagaatag	agaacttgct	ttggtccaga	caaaaacaat	taaaaagaca	600
ctgaacccaa	aatggaatga	agaattttat	ttcagggtaa	acccatctaa	tcacagactc	660
ctatttgaag	tatttgacga	aaatagactg	acacgagacg	acttcctggg	ccaggtggac	720
gtgcccctta	gtcaccttcc	gacagaagat	ccaaccatgg	agcgacccta	tacatttaag	780
gactttctcc	tcagaccaag	aagtcataag	tctcgagtta	agggattttt	gcgattgaaa	840
atggcctata	tgccaaaaaa	tggaggtcaa	gatgaagaaa	acagtgacca	gagggatgac	900
atggagcatg	gatgggaagt	tgttgactca	aatgactcgg	cttctcagca	ccaagaggaa	960
cttctctctc	ctctcttgcc	tcccgggtgg	gaagaaaaag	tggacaattt	aggccgaact	1020
tactatgtca	accacaacaa	ccggaccact	cagtggcaca	gaccaagcct	gatggacgtg	1080
tcctcggagt	cggacaataa	catcagacag	atcaaccagg	aggcagcaca	ccggcgcttc	1140
cgctcccgcg	ggcacatcag	cgaagacttg	gagcccagac	cctcggaggg	cggggatgtc	1200
cccgagcctt	gggagaccat	ttcagaggaa	gtgaatatcg	ctggagactc	tctcgggtgtg	1260
gttttgcccc	caccaccggc	ctccccagga	tctcggacca	gccctcagga	gctgtcagag	1320
gaactaagca	gaaggcttca	gatcactcca	gactccaatg	gggaacagtt	cagctctttg	1380
attcaaagag	aaccctcttc	aagggttgagg	tcatgcagtg	tcaccgacgc	agttgcagaa	1440
cagggccatc	taccaccgcc	atcagtggcc	tatgtacata	ccacgccggg	tctgccttca	1500
ggctgggaag	aaagaaaaga	tgctaagggg	cgcacatact	atgtcaatca	taacaatcga	1560
accacaactt	ggactcgacc	tatcatgcag	cttgccagaag	atgggtgcgtc	cggatcagcc	1620
acaaacagta	acaaccatct	aatcgagcct	cagatccgcc	ggcctcgtag	cctcagctcg	1680
ccaacagtaa	ctttattgcc	ccgctggagg	gtgccaagga	ctcaccgcga	cgtcgggctg	1740



tgaaagacac	cctttccaac	ccacagtcce	cacagccatc	accttacaac	ccccccaaac	1800
cacaacacaa	agtcacacag	agcttcttgc	caccgcgctg	ggaaatgagg	atagcgccaa	1860
acggccggcc	cttcttcatt	gatacataca	caaagactac	aacctgggaa	gatccacgtt	1920
tgaaatttcc	agtacatatg	cggtaaaaga	catctttaaa	ccccaatgac	cttggccccc	1980
ttcctcctgg	ctgggaagaa	agaattcact	tggatggccg	aacgttttat	attgatcata	2040
atagcaaaat	tactcagtgg	gaagacccaa	gactgcagaa	cccagctatt	actggtcagg	2100
ctgtccctta	ctccagagaa	tttaagcaga	aatatgacta	cttcaggaag	aaatttaaaga	2160
aacctgctga	tatccccaat	aggtttgaaa	tgaaacttca	cagaaataac	atatttgaag	2220
agtcctatcg	gagaatttatg	tcogtgaaaa	gaccagatgt	cctaaaagct	agactgtgga	2280
ttgagtttga	atcagagaaa	ggctctgact	atgggggtgt	ggccagagaa	tggttcttct	2340
tactgtccaa	agagatgttc	aacccttact	acggcctctt	tgagtactct	gccacggaca	2400
actacacctt	tcagatcaac	cctaattcag	gcctctgtaa	tgaggatcat	ttgtcctact	2460
tcacttttat	tggaagagtt	gctggctctg	cogtatttca	tggaagctc	ttagatgggt	2520
tcttcattag	accattttac	aagatgatgt	tgggaaagca	gataaccctg	aatgacatgg	2580
aatctgtgga	tagtgaatat	tacaactctt	tgaaatggat	cctggagaat	gacctactg	2640
agctggacct	catgttctgc	atagacgaag	aaaactttgg	acagacatat	caagtggatt	2700
tgaagcccaa	tgggtcagaa	ataatggtca	caaatgaaaa	caaaagggaa	tatatcgact	2760
tagtcatcca	gtggagattt	gtgaacaggg	tccagaagca	gatgaacgcc	ttcttggagg	2820
gattcacaga	actacttcc	attgatttga	ttaaaatttt	tgatgaaaat	gagctggagt	2880
tgctcatgtg	cggcctcggt	gatgtggatg	tgaatgactg	gagacagcat	tctatttaca	2940
agaacggcta	ctgccccaac	caccccgta	ttcagtggtt	ctggaaggct	gtgctactca	3000
tggacgccga	aaagegtatc	cggttactgc	agtttgacac	agggacatcg	cgagtaccta	3060
tgaatggatt	tgcgaactt	tatggttcca	atggtcctca	gctgtttaca	atagagcaat	3120
ggggcagtc	tgagaaaactg	cccagagctc	acacatgctt	taatcgccct	gacttacctc	3180
catatgaaac	ctttgaagat	ttacgagaga	aacttctcat	ggccgtggaa	aatgctcaag	3240
gatttgaagg	ggtggattaa	gcacctgtg	cctcggggtg	ggttggttct	caagcaagtt	3300
ctgcttgcac	ttttgcattt	gcctaacaga	cttttgacaga	ggcgatggca	gagagcagct	3360
gcaggcatgg	tccttggaagc	cgagccttca	ccacgcactc	gtccaagttc	gggatgcggg	3420
aacctgggtc	cagcttgagt	tctgccttt	cccaccacaa	attatcaact	ggttgatgtg	3480
tacactaatt	acatttcagg	aggacttaat	gctatttatg	ttgtcctctg	caggcaaaagc	3540
ccttaataaa	tattttacat	cctttctaat	gacaatgaat	ggaattaatc	actcaacagg	3600
tatagtatta	cgactcatgt	ttacttttta	aaatgattta	gaccgatttt	cagattttat	3660
ttcgttatga	ttaaagatgt	ctcatgtact	tggaaaagtg	agcatttttt	tttttttttg	3720
tatttcactt	tcataaccagg	cctaattgtca	atgacatttt	tatttttgaa	gtactctgac	3780
acctccaccc	tctactttat	tagaattgga	aggcaaat	ttgtccaaaa	acctacagac	3840
aagtactttg	agagaatttc	caatataata	ttagacataa	tgataatttt	ttccatactc	3900
agaatgaaaa	actggatatt	acgtttttgt	tttgggggtt	ttttgtacaa	atttagctaa	3960
tagctacagg	ctgagagaa	tgtaacatag	catgacaaat	tttgtgttga	cttgaaagga	4020
atcacaccat	tattccttag	aagtaattac	atgtgttcta	acacatttga	gacaggggtg	4080
gactcccatt	tctcatccga	gaaattactt	aaccttctct	gggcgctgta	cagtcactct	4140
ttattctatt	tcctctttgc	tgtttgtagt	agagacattt	tgaatgaaac	ttggcactgc	4200
ttgattcaaa	actgtggaaa	ccagatctgt	ttagtctcct	gtttgatatg	gtttgcta	4260
ggtagctaaa	taaccagttt	ttgttgtaaa	tgcaccaatt	ctgaaggcac	tttatgtact	4320
acatggaggt	catatctggg	tttgttttta	tttttttata	atgaacatta	aatgtgatga	4380
tgatttcttt	tcctgcaca	catctttccg	gtgcaatata	tatcaattgt	gaatctggct	4440
gctgggtgat	aaaaacctgg	atgtaaagct	gagcctacag	acctgtcctc	accaactgtt	4500
ttgtgatttc	tactcaacta	caaagattta	tttaatgtac	tcttaatcta	actgagtttt	4560
gttaccaatg	acctgtttga	tgtttcaata	ccgtgtactg	cctgagttgt	gcctcttggtg	4620
tgctagatta	aaagtggagc	agagacttga	cttgatcctc	tgagcctcaa	gctattgagc	4680
tggtagtggc	agaggactga	gggtacctgc	acagtttgat	tcttttccca	cgttgtaagt	4740
ctccattgca	gaattgtcgt	gcgtttgaga	aaacacctga	ggcagtggtg	gagttgaacg	4800
acctgctgtg	cctttttaac	ctgtgttgct	ctagacctgt	cggggcagtc	aggggacact	4860
agagatttga	tctcatgcga	gtcatcaata	ggacaaaaaa	gttggtgggtt	ggggaggtct	4920
gtttgttaca	taaaaaggac	ctttcggtgt	aagaaattgc	cgtttttacc	ctgccctggc	4980
tggcatgtga	gaagccatgg	aagggtgtgg	ttgtaaatga	gttgtctaaa	gggggtgcaga	5040
ggcctgaggt	ttctaaaaga	aggtagattt	ctacagagct	gagtggttgt	tcctttttct	5100
tattggttga	aaattacctg	gtagtgatca	gaaaacttag	atgctatgta	actc	5154

<210> 2
 <211> 975
 <212> PRT
 <213> Homo sapiens

<220>
 <221> MOD_RES
 <222> (455)
 <223> Any amino acid

<400> 2
 Met Phe Arg Leu Arg Ser Trp Ala Ser Ser Thr Thr Gly Ser Arg Tyr
 1 5 10 15
 Gly Ser Ala Phe Cys Gly Ser Pro Thr Leu Ala Trp Cys Val Cys Val
 20 25 30
 Pro Val Cys Tyr Gly Glu Ser Arg Ile Leu Arg Val Lys Val Val Ser
 35 40 45
 Gly Ile Asp Leu Ala Lys Lys Asp Ile Phe Gly Ala Ser Asp Pro Tyr
 50 55 60
 Val Lys Leu Ser Leu Tyr Val Ala Asp Glu Asn Arg Glu Leu Ala Leu
 65 70 75 80
 Val Gln Thr Lys Thr Ile Lys Lys Thr Leu Asn Pro Lys Trp Asn Glu
 85 90 95
 Glu Phe Tyr Phe Arg Val Asn Pro Ser Asn His Arg Leu Leu Phe Glu
 100 105 110
 Val Phe Asp Glu Asn Arg Leu Thr Arg Asp Asp Phe Leu Gly Gln Val
 115 120 125
 Asp Val Pro Leu Ser His Leu Pro Thr Glu Asp Pro Thr Met Glu Arg
 130 135 140
 Pro Tyr Thr Phe Lys Asp Phe Leu Leu Arg Pro Arg Ser His Lys Ser
 145 150 155 160
 Arg Val Lys Gly Phe Leu Arg Leu Lys Met Ala Tyr Met Pro Lys Asn
 165 170 175
 Gly Gly Gln Asp Glu Glu Asn Ser Asp Gln Arg Asp Asp Met Glu His
 180 185 190
 Gly Trp Glu Val Val Asp Ser Asn Asp Ser Ala Ser Gln His Gln Glu
 195 200 205
 Glu Leu Pro Pro Pro Pro Leu Pro Pro Gly Trp Glu Glu Lys Val Asp
 210 215 220
 Asn Leu Gly Arg Thr Tyr Tyr Val Asn His Asn Asn Arg Thr Thr Gln
 225 230 235 240

Trp His Arg Pro Ser Leu Met Asp Val Ser Ser Glu Ser Asp Asn Asn
 245 250 255
 Ile Arg Gln Ile Asn Gln Glu Ala Ala His Arg Arg Phe Arg Ser Arg
 260 265 270
 Arg His Ile Ser Glu Asp Leu Glu Pro Glu Pro Ser Glu Gly Gly Asp
 275 280 285
 Val Pro Glu Pro Trp Glu Thr Ile Ser Glu Glu Val Asn Ile Ala Gly
 290 295 300
 Asp Ser Leu Gly Val Val Leu Pro Pro Pro Pro Ala Ser Pro Gly Ser
 305 310 315 320
 Arg Thr Ser Pro Gln Glu Leu Ser Glu Glu Leu Ser Arg Arg Leu Gln
 325 330 335
 Ile Thr Pro Asp Ser Asn Gly Glu Gln Phe Ser Ser Leu Ile Gln Arg
 340 345 350
 Glu Pro Ser Ser Arg Leu Arg Ser Cys Ser Val Thr Asp Ala Val Ala
 355 360 365
 Glu Gln Gly His Leu Pro Pro Pro Ser Val Ala Tyr Val His Thr Thr
 370 375 380
 Pro Gly Leu Pro Ser Gly Trp Glu Glu Arg Lys Asp Ala Lys Gly Arg
 385 390 395 400
 Thr Tyr Tyr Val Asn His Asn Asn Arg Thr Thr Thr Trp Thr Arg Pro
 405 410 415
 Ile Met Gln Leu Ala Glu Asp Gly Ala Ser Gly Ser Ala Thr Asn Ser
 420 425 430
 Asn Asn His Leu Ile Glu Pro Gln Ile Arg Arg Pro Arg Ser Leu Ser
 435 440 445
 Ser Pro Thr Val Thr Leu Xaa Ala Pro Leu Glu Gly Ala Lys Asp Ser
 450 455 460
 Pro Val Arg Arg Ala Val Lys Asp Thr Leu Ser Asn Pro Gln Ser Pro
 465 470 475 480
 Gln Pro Ser Pro Tyr Asn Ser Pro Lys Pro Gln His Lys Val Thr Gln
 485 490 495
 Ser Phe Leu Pro Pro Gly Trp Glu Met Arg Ile Ala Pro Asn Gly Arg
 500 505 510
 Pro Phe Phe Ile Asp His Asn Thr Lys Thr Thr Thr Trp Glu Asp Pro
 515 520 525
 Arg Leu Lys Phe Pro Val His Met Arg Ser Lys Thr Ser Leu Asn Pro
 530 535 540

Asn Asp Leu Gly Pro Leu Pro Pro Gly Trp Glu Glu Arg Ile His Leu
 545 550 555 560
 Asp Gly Arg Thr Phe Tyr Ile Asp His Asn Ser Lys Ile Thr Gln Trp
 565 570 575
 Glu Asp Pro Arg Leu Gln Asn Pro Ala Ile Thr Gly Pro Ala Val Pro
 580 585 590
 Tyr Ser Arg Glu Phe Lys Gln Lys Tyr Asp Tyr Phe Arg Lys Lys Leu
 595 600 605
 Lys Lys Pro Ala Asp Ile Pro Asn Arg Phe Glu Met Lys Leu His Arg
 610 615 620
 Asn Asn Ile Phe Glu Glu Ser Tyr Arg Arg Ile Met Ser Val Lys Arg
 625 630 635 640
 Pro Asp Val Leu Lys Ala Arg Leu Trp Ile Glu Phe Glu Ser Glu Lys
 645 650 655
 Gly Leu Asp Tyr Gly Gly Val Ala Arg Glu Trp Phe Phe Leu Leu Ser
 660 665 670
 Lys Glu Met Phe Asn Pro Tyr Tyr Gly Leu Phe Glu Tyr Ser Ala Thr
 675 680 685
 Asp Asn Tyr Thr Leu Gln Ile Asn Pro Asn Ser Gly Leu Cys Asn Glu
 690 695 700
 Asp His Leu Ser Tyr Phe Thr Phe Ile Gly Arg Val Ala Gly Leu Ala
 705 710 715 720
 Val Phe His Gly Lys Leu Leu Asp Gly Phe Phe Ile Arg Pro Phe Tyr
 725 730 735
 Lys Met Met Leu Gly Lys Gln Ile Thr Leu Asn Asp Met Glu Ser Val
 740 745 750
 Asp Ser Glu Tyr Tyr Asn Ser Leu Lys Trp Ile Leu Glu Asn Asp Pro
 755 760 765
 Thr Glu Leu Asp Leu Met Phe Cys Ile Asp Glu Glu Asn Phe Gly Gln
 770 775 780
 Thr Tyr Gln Val Asp Leu Lys Pro Asn Gly Ser Glu Ile Met Val Thr
 785 790 795 800
 Asn Glu Asn Lys Arg Glu Tyr Ile Asp Leu Val Ile Gln Trp Arg Phe
 805 810 815
 Val Asn Arg Val Gln Lys Gln Met Asn Ala Phe Leu Glu Gly Phe Thr
 820 825 830
 Glu Leu Leu Pro Ile Asp Leu Ile Lys Ile Phe Asp Glu Asn Glu Leu
 835 840 845

Glu Leu Leu Met Cys Gly Leu Gly Asp Val Asp Val Asn Asp Trp Arg
 850 855 860

Gln His Ser Ile Tyr Lys Asn Gly Tyr Cys Pro Asn His Pro Val Ile
 865 870 875 880

Gln Trp Phe Trp Lys Ala Val Leu Leu Met Asp Ala Glu Lys Arg Ile
 885 890 895

Arg Leu Leu Gln Phe Val Thr Gly Thr Ser Arg Val Pro Met Asn Gly
 900 905 910

Phe Ala Glu Leu Tyr Gly Ser Asn Gly Pro Gln Leu Phe Thr Ile Glu
 915 920 925

Gln Trp Gly Ser Pro Glu Lys Leu Pro Arg Ala His Thr Cys Phe Asn
 930 935 940

Arg Leu Asp Leu Pro Pro Tyr Glu Thr Phe Glu Asp Leu Arg Glu Lys
 945 950 955 960

Leu Leu Met Ala Val Glu Asn Ala Gln Gly Phe Glu Gly Val Asp
 965 970 975

<210> 3
 <211> 854
 <212> DNA
 <213> Murine sp.

<400> 3
 acaatggggg cgtggcagag aatggttctt cttactgtcc aaagagatgt ttaacccta 60
 ctatggcctc ttcgagtact ctgccacgga caactacaca cttcagatca atcccaactc 120
 aggcctctgt aatgaagacc atttgtccta tttcaccttc attggaagag ttgctggcct 180
 agcgggtgtt catgggaaac tcttagatgg attcttcatt cgaccattct acaagatgat 240
 gctggggaag cagataacgc tgaacgacat ggagtccgtg gacagcgagt actacaactc 300
 tttgaagtgg atcttagaaa acgacccac ggaacttgac ctcattgttct gcatagacga 360
 gagaactttg ggcagacata ccaagtggat ctgaagccca acgggtcaga aataatggta 420
 accaatgaga acaaacgaga atacattgac ttagtcatcc agtggagatt tgtgaacagg 480
 gtccagaagc aaatgaatgc cttcttggag ggatttacag aacttcttcc aatcgacttg 540
 attaaaattt ttgatgaaaa tgagctggag ttgctgatgt gcggccttgg tgatgtcgac 600
 gtgaacgact ggagacagca ctctatttac aagaacggct actgccccaa ccaccctgtc 660
 atccagtgg tctggaaggc cgtgctcctg atggatgctg agaagcgcat ccggttacta 720
 cagtttgtca caggcacctc cagagtaccc atgaatggat ttgccgaact ctatggttcc 780
 aatggtcctc agctgtttac aatagagcaa tggggcagtc cgaaaaacta ccagagctct 840
 acatgcttaa tcgc 854

<210> 4
 <211> 604
 <212> PRT
 <213> Murine sp.

<220>
 <221> MOD_RES
 <222> (561)
 <223> Any amino acid

<400> 4

His	Ala	Cys	Ser	Asn	Ala	Ala	Ser	Arg	Ala	Ala	Ala	Arg	Val	Ala	Ala		
1				5					10					15			
Arg	Cys	Thr	Ala	Arg	Ser	Arg	Ser	Gly	Arg	Arg	Ser	Ser	Ser	Val	Ser		
		20						25					30				
Arg	Ser	Ser	Ser	Arg	Gly	Ala	Ser	Ser	Ser	Met	Ser	Ser	Asp	Met	Ala		
		35					40					45					
Ala	Asp	Ser	Ala	Val	Ser	Asp	Val	Trp	Cys	Asp	Lys	Thr	Asp	Gly	Gly		
	50					55					60						
Gly	Ser	Gly	Ser	Asp	Val	Thr	Asp	Thr	Cys	Cys	Gly	Cys	Trp	Asn	Asn		
65					70					75					80		
Ser	His	Val	Thr	Ala	Asp	Tyr	His	Asn	Asp	Asp	Thr	Arg	Val	Val	Arg		
				85				90						95			
Val	Lys	Val	Ala	Gly	Gly	Ala	Lys	Lys	Asp	Gly	Ala	Ser	Asp	Tyr	Val		
		100						105					110				
Arg	Val	Thr	Tyr	Asp	Met	Ser	Gly	Thr	Ser	Val	Thr	Lys	Thr	Lys	Lys		
		115					120					125					
Ser	Asn	Lys	Trp	Asn	Arg	Val	Arg	His	Arg	Val	Asp	Asn	Arg	Thr	Arg		
	130					135					140						
Asp	Asp	Gly	Val	Asp	Val	Tyr	Thr	Asn	Arg	Met	Arg	Tyr	Thr	Lys	Asp		
145					150					155				160			
Val	His	Arg	Ser	His	Lys	Ser	Arg	Val	Lys	Gly	Tyr	Arg	Lys	Met	Thr		
				165					170					175			
Tyr	Lys	Asn	Gly	Ser	Asp	Asn	Ala	Asp	Ala	Gly	Trp	Val	Val	Asp	Asp		
		180						185					190				
Ala	Ala	Thr	His	His	Ser	Gly	Trp	Arg	Asp	Val	Gly	Arg	Thr	Tyr	Tyr		
		195					200					205					
Val	Asn	His	Ser	Arg	Arg	Thr	Trp	Lys	Arg	Ser	Asp	Asp	Asp	Thr	Asp		
	210					215					220						
Asp	Asn	Asp	Asp	Met	Ala	Arg	Ala	Thr	Thr	Arg	Arg	Ser	Asp	Val	Asp		
225					230					235				240			
Gly	Asp	Asn	Arg	Ser	Asn	Trp	Val	Arg	Asp	Asn	Thr	Tyr	Ser	Gly	Ala		
				245					250					255			
Val	Ser	Ser	Gly	His	Asp	Val	Thr	His	Ala	Asn	Thr	Arg	Ala	Val	Cys		
			260					265					270				
Gly	Asn	Ala	Thr	Ser	Val	Thr	Ser	Ser	Asn	His	Ser	Ser	Arg	Gly	Gly		
	275						280					285					
Ser	Thr	Cys	Thr	Val	Thr	Ser	Ser	Gly	Gly	Trp	Lys	Asp	Asp	Arg	Gly		
	290					295					300						

Arg Ser Tyr Tyr Val Asp His Asn Ser Lys Thr Thr Thr Trp Ser Lys
 305 310 315 320
 Thr Met Asp Asp Arg Ser Lys Ala His Arg Gly Lys Thr Asp Ser Asn
 325 330 335
 Asp Gly Gly Trp Arg Thr His Thr Asp Gly Arg Val Asn His Asn Lys
 340 345 350
 Lys Thr Trp Asp Arg Asn Val Ala Thr Gly Ala Val Tyr Ser Arg Asp
 355 360 365
 Tyr Lys Arg Lys Tyr Arg Arg Lys Lys Lys Thr Asp Asn Lys Met Lys
 370 375 380
 Arg Arg Ala Asn Asp Ser Tyr Arg Arg Met Gly Val Lys Arg Ala Asp
 385 390 395 400
 Lys Ala Arg Trp Asp Gly Lys Gly Asp Tyr Gly Gly Val Ala Arg Trp
 405 410 415
 Ser Lys Met Asn Tyr Tyr Gly Tyr Ser Ala Thr Asp Asn Tyr Thr Asn
 420 425 430
 Asn Ser Gly Cys Asn Asp His Ser Tyr Lys Gly Arg Val Ala Gly Met
 435 440 445
 Ala Val Tyr His Gly Lys Asp Gly Arg Tyr Lys Met Met Lys Thr His
 450 455 460
 Asp Met Ser Val Asp Ser Tyr Tyr Ser Ser Arg Trp Asn Asp Thr Asp
 465 470 475 480
 Arg Asp Gly Thr His His Lys Thr Gly Gly Ser Val Val Thr Asn Lys
 485 490 495
 Asn Lys Lys Tyr Tyr Val Trp Arg Val Asn Arg Lys Met Ala Ala Lys
 500 505 510
 Gly Asp Lys Asp Asn Met Cys Gly Gly Asp Val Asp Val Asn Asp Trp
 515 520 525
 Arg His Thr Lys Tyr Lys Asn Gly Tyr Ser Met Asn His Val His Trp
 530 535 540
 Trp Lys Ala Val Trp Met Met Asp Ser Lys Arg Arg Val Thr Gly Thr
 545 550 555 560
 Xaa Ser Arg Val Met Asn Gly Ala Tyr Gly Ser Asn Gly Ser Thr Val
 565 570 575
 Trp Gly Thr Asp Lys Arg Ala His Thr Cys Asn Arg Asp Tyr Ser Asp
 580 585 590
 Trp Asp Lys Met Ala Asn Thr Gly Asp Gly Val Asp
 595 600

<210> 5
 <211> 615
 <212> DNA
 <213> Homo sapiens

<400> 5
 tgctgcaagt gacaggttcc aagaagcccc agggctcaga gctgaatgat gaagcgcagt 60
 ccccaaagtg cctggccacc cctccctccc tggatcactg ctgcctgggc ttgattgatt 120
 gattgattga ttgattgatt gattttgaga gagattctca ctgtcaccca ggctggagta 180
 cagtgggtgcg atctcggctc actgcagcct ctgcctcccc gggttcaagca attctcctgc 240
 ctcagcctcc caagtagctg ggactacagg cacgcgccac cacacccagc taatttttcta 300
 ttttttagtaa aagacggggt ttcaccatgt tggggcagga tggctcttgat ctcttgacct 360
 catgatccac ccgccccggc ttccaaagtg ctgggataca ggcataaacc cgacgcgccc 420
 agcatggaca ttttttttta atccccctgcc cttttcttgg gcataattca ttgcaggctc 480
 cttctataca gatcatggaa aacacatttt cttaactgag ttttattatt tatacccagc 540
 acctcatgac atttaccctg ttacaacaaa atgggcacct gccaaaacaa ctttatataa 600
 ggatgctcca ggcct 615

<210> 6
 <211> 927
 <212> PRT
 <213> Homo sapiens

<400> 6
 Ser Arg Phe Ser Ser Ser Ser Thr Val Ala Cys Pro Gly Arg Gly
 1 5 10 15
 Arg Ala Arg Pro Val Cys Trp Lys Arg Ser Glu Met Ala Thr Cys Ala
 20 25 30
 Val Glu Val Phe Gly Leu Leu Glu Asp Glu Glu Asn Ser Arg Ile Val
 35 40 45
 Arg Val Arg Val Ile Ala Gly Ile Gly Leu Ala Lys Lys Asp Ile Leu
 50 55 60
 Gly Ala Ser Asp Pro Tyr Val Arg Val Thr Leu Tyr Asp Pro Met Asn
 65 70 75 80
 Gly Val Leu Thr Ser Val Gln Thr Lys Thr Ile Lys Lys Ser Leu Asn
 85 90 95
 Pro Lys Trp Asn Glu Glu Ile Leu Phe Arg Val His Pro Gln Gln His
 100 105 110
 Arg Leu Leu Phe Glu Val Phe Asp Glu Asn Arg Leu Thr Arg Asp Asp
 115 120 125
 Phe Leu Gly Gln Val Asp Val Pro Leu Tyr Pro Leu Pro Thr Glu Asn
 130 135 140
 Pro Arg Leu Glu Arg Pro Tyr Thr Phe Lys Asp Phe Val Leu His Pro
 145 150 155 160
 Arg Ser His Lys Ser Arg Val Lys Gly Tyr Leu Arg Leu Lys Met Thr
 165 170 175

Tyr Leu Pro Lys Thr Ser Gly Ser Glu Asp Asp Asn Ala Glu Gln Ala
 180 185 190
 Glu Glu Leu Glu Pro Gly Trp Val Val Leu Asp Gln Pro Asp Ala Ala
 195 200 205
 Cys His Leu Gln Gln Gln Gln Glu Pro Ser Pro Leu Pro Pro Gly Trp
 210 215 220
 Glu Glu Arg Gln Asp Ile Leu Gly Arg Thr Tyr Tyr Val Asn His Glu
 225 230 235 240
 Ser Arg Arg Thr Gln Trp Lys Arg Pro Thr Pro Gln Asp Asn Leu Thr
 245 250 255
 Asp Ala Glu Asn Gly Asn Ile Gln Leu Gln Ala Gln Arg Ala Phe Thr
 260 265 270
 Thr Arg Arg Gln Ile Ser Glu Glu Thr Glu Ser Val Asp Asn Gln Glu
 275 280 285
 Ser Ser Glu Asn Trp Glu Ile Ile Arg Glu Asp Glu Ala Thr Met Tyr
 290 295 300
 Ser Ser Gln Ala Phe Pro Ser Pro Pro Pro Ser Ser Asn Leu Asp Val
 305 310 315 320
 Pro Thr His Leu Ala Glu Glu Leu Asn Ala Arg Leu Thr Ile Phe Gly
 325 330 335
 Asn Ser Ala Val Ser Gln Pro Ala Ser Ser Ser Asn His Ser Ser Arg
 340 345 350
 Arg Gly Ser Leu Gln Ala Tyr Thr Phe Glu Glu Gln Pro Thr Leu Pro
 355 360 365
 Val Leu Leu Pro Thr Ser Ser Gly Leu Pro Pro Gly Trp Glu Glu Lys
 370 375 380
 Gln Asp Glu Arg Gly Arg Ser Tyr Tyr Val Asp His Asn Ser Arg Thr
 385 390 395 400
 Thr Thr Trp Thr Lys Pro Thr Val Gln Ala Thr Val Glu Thr Ser Gln
 405 410 415
 Leu Thr Ser Ser Gln Ser Ser Ala Gly Pro Gln Ser Gln Ala Ser Thr
 420 425 430
 Ser Asp Ser Gly Gln Gln Val Thr Gln Pro Ser Glu Ile Glu Gln Gly
 435 440 445
 Phe Leu Pro Lys Gly Trp Glu Val Arg His Ala Pro Asn Gly Arg Pro
 450 455 460
 Phe Phe Ile Asp His Asn Thr Lys Thr Thr Thr Trp Glu Asp Pro Arg
 465 470 475 480

Leu	Lys	Ile	Pro	Ala	His	Leu	Arg	Gly	Lys	Thr	Ser	Leu	Asp	Thr	Ser	485	490	495
Asn	Asp	Leu	Gly	Pro	Leu	Pro	Pro	Gly	Trp	Glu	Glu	Arg	Thr	His	Thr	500	505	510
Asp	Gly	Arg	Ile	Phe	Tyr	Ile	Asn	His	Asn	Ile	Lys	Arg	Thr	Gln	Trp	515	520	525
Glu	Asp	Pro	Arg	Leu	Glu	Asn	Val	Ala	Ile	Thr	Gly	Pro	Ala	Val	Pro	530	535	540
Tyr	Ser	Arg	Asp	Tyr	Lys	Arg	Lys	Tyr	Glu	Phe	Phe	Arg	Arg	Lys	Leu	545	550	555
Lys	Lys	Gln	Asn	Asp	Ile	Pro	Asn	Lys	Phe	Glu	Met	Lys	Leu	Arg	Arg	565	570	575
Ala	Thr	Val	Leu	Glu	Asp	Ser	Tyr	Arg	Arg	Ile	Met	Gly	Val	Lys	Arg	580	585	590
Ala	Asp	Phe	Leu	Lys	Ala	Arg	Leu	Trp	Ile	Glu	Phe	Asp	Gly	Glu	Lys	595	600	605
Gly	Leu	Asp	Tyr	Gly	Gly	Val	Ala	Arg	Glu	Trp	Phe	Phe	Leu	Ile	Ser	610	615	620
Lys	Glu	Met	Phe	Asn	Pro	Tyr	Tyr	Gly	Leu	Phe	Glu	Tyr	Ser	Ala	Thr	625	630	635
Asp	Asn	Tyr	Thr	Leu	Gln	Ile	Asn	Pro	Asn	Ser	Gly	Leu	Cys	Asn	Glu	645	650	655
Asp	His	Leu	Ser	Tyr	Phe	Lys	Phe	Ile	Gly	Arg	Val	Ala	Gly	Met	Ala	660	665	670
Val	Tyr	His	Gly	Lys	Leu	Leu	Asp	Gly	Phe	Phe	Ile	Arg	Pro	Phe	Tyr	675	680	685
Lys	Met	Met	Leu	His	Lys	Pro	Ile	Thr	Leu	His	Asp	Met	Glu	Ser	Val	690	695	700
Asp	Ser	Glu	Tyr	Tyr	Asn	Ser	Leu	Arg	Trp	Ile	Leu	Glu	Asn	Asp	Pro	705	710	715
Thr	Glu	Leu	Asp	Leu	Arg	Phe	Ile	Ile	Asp	Glu	Glu	Leu	Phe	Gly	Gln	725	730	735
Thr	His	Gln	His	Glu	Leu	Lys	Asn	Gly	Gly	Ser	Glu	Ile	Val	Val	Thr	740	745	750
Asn	Lys	Asn	Lys	Lys	Glu	Tyr	Ile	Tyr	Leu	Val	Ile	Gln	Trp	Arg	Phe	755	760	765
Val	Asn	Arg	Ile	Gln	Lys	Gln	Met	Ala	Ala	Phe	Lys	Glu	Gly	Phe	Phe	770	775	780

Glu Leu Ile Pro Gln Asp Leu Ile Lys Ile Phe Asp Glu Asn Glu Leu
785 790 795 800

Glu Leu Leu Met Cys Gly Leu Gly Asp Val Asp Val Asn Asp Trp Arg
805 810 815

Glu His Thr Lys Tyr Lys Asn Gly Tyr Ser Ala Asn His Gln Val Ile
820 825 830

Gln Trp Phe Trp Lys Ala Val Leu Met Met Asp Ser Glu Lys Arg Ile
835 840 845

Arg Leu Leu Gln Phe Val Thr Gly Thr Ser Arg Val Pro Met Asn Gly
850 855 860

Phe Ala Glu Leu Tyr Gly Ser Asn Gly Pro Gln Ser Phe Thr Val Glu
865 870 875 880

Gln Trp Gly Thr Pro Glu Lys Leu Pro Arg Ala His Thr Cys Phe Asn
885 890 895

Arg Leu Asp Leu Pro Pro Tyr Glu Ser Phe Glu Glu Leu Trp Asp Lys
900 905 910

Leu Gln Met Ala Ile Glu Asn Thr Gln Gly Phe Asp Gly Val Asp
915 920 925

<210> 7
<211> 879
<212> DNA
<213> Homo sapiens

<400> 7
cagagaaagg tcttgactat ggggggtgtgg ccagagaatg gttctttctta ctctccaaag 60
agatgttcaa cccctactac ggctcttttg agtactctgc caccacacaac tacacccttc 120
agatcaaccc taattcaggc ctctgtaatc acgatcattt gtctacttc acttttattg 180
gaagagttgc tgggtctggcc gtatttcatg ggaagctctt agatggtttc ttcattagac 240
cattttacaa gatgatgttg ggaaagcaga taaccttgaa tcacatggaa tctgtggata 300
gtgaatatta caactctttg aaatggatcc tccacaatga ccctactgag ctggacctca 360
tgttctgcat agacgaagaa aactttggac agacatatca agtggatttg aagcccaatg 420
ggtcacaaat aatgctcaca aatgaaaaca aaaggggaata tatcgactta gtcattccagt 480
ggagatttgt gaacagggtc cagaagcaga tgaacgcctt cttggaggga ttcacagaac 540
tacttctat tgatttgatt aaaatttttg atgaaaatga gctggagttg ctcatgtgcg 600
gcctcgggtga tgtggatgtg aatgactgga gacagcattc tatttacaag aacggctact 660
gcccaaacca ccccgtcatt cagtggttct ggaaggctgt gctactcatg gacgccgaaa 720
agcgtatccg gttactgcag tttgtcacag ggacatcccg agtacctatg aatggatttg 780
ccgaacttta tggttccaat ggctctcagc tgtttacaat agagcaatgg ggcagtctg 840
agaaactgcc cagagctcac acatgcttta atcgcttg 879

<210> 8
<211> 3060
<212> DNA
<213> Homo sapiens

<400> 8

```

cggatatcagc agaggtgtgt acggggcactg ctttaaaact gggaaggagg aagacgaggc 60
cagggagccg gaggggtcac aaggttagatt tccagcagcg ctagtccagc tgaacacttt 120
ccagccttgt ttttcagcag ctttgaggaa aagtatagt atccgtatgt gaaactttca 180
ttgtacgtag cggatgagaa tagagaactt gctttggtcc agacaaaaac aattaaaaag 240
acactgaacc caaaatggaa tgaagaattt tatttcaggg taaacccatc taatcacaga 300
ctcctatttg aagtatttga cgaaaataga ctgacacgag acggcttctc gggccaggtg 360
gacgtgcccc ttagtcacct tccgacagaa gatccaacca tggagcgacc ctatacattt 420
aaggactttc tcctcagacc aagaagtcac aagtctcgag ttaagggatt tttgcgattg 480
aaaatggcct atatgccaaa aaatggaggt caagatgaag aaaacagtga ccagagggat 540
gacatggagc atggatggga agttgttgac tcaaatgact cggcttctca gcaccaagag 600
gaacttctc tcctcctctc gctcccggg tgggaagaaa aagtggacaa tttaggccga 660
acttactatg tcaaccacaa caaccggacc actcagtggc acagaccaag cctgatggac 720
gtgtcctcgg agtcggacaa taacatcaga cagatcaacc aggaggcagc acaccggcgc 780
ttccgctccc gcaggcacat cagcgaagac ttggagcccc agcctcggga gggcggggat 840
gtccccgagc cttgggagac catttcagag gaagtgaata tcgctggaga ctctctcggt 900
ctggctctgc ccccaccacc ggtctcccca ggcctcggga ccagccctca ggactgtca 960
gaggaactaa gcagaaggct tcagatcact ccagactcca atggggaaca gttcagctct 1020
ttgattcaaa gagaacctc ctcaagggtg aggtcatgca gtgtcacga cgcagttgca 1080
gaacagggcc atctaccacc gccatcagtg gcctatgtac ataccacgcc gggctctgct 1140
tcaggctggg aagaaagaaa agatgctaag gggcgacat actatgtcaa tcataacaat 1200
cgaaccacaa cttggactcg acctatcatg cagcttgtag aagatgggtg gtccggatca 1260
gccacaaaca gtaacaacca tctaatecag cctcagatcc gccggcctcg tagcctcagc 1320
tcgccaacag taactttatc tgccccgctg gaggggtgcca aggactcacc cgtacgtcgg 1380
gctgtgaaag acaccctttc caaccacag tccccacagc catcacctta caactcccc 1440
aaaccacaac acaaagtcac acagagcttc ttgccaccg gctgggaaat gaggatagcg 1500
ccaaacggcc ggcccttctt cattgatcat aacacaaaga caacaacctg ggaagatcca 1560
cgtttgaaat ttccagtaca tatgcggtca aagacatctt taaaccccaa tgaccttggc 1620
ccccttctc ctggctggga agaaagaatt cacttggatg gccgaacgtt ttatattgat 1680
cataatagca aaattactca gtgggaagac ccaagactgc agaaccacgc tattactggt 1740
ccggctgtcc ctactccag agaatttaag cagaaatatg actacttcag gaagaaatta 1800
aagaaacctg ctgatatccc caatagggtt gaaatgaaac ttcacagaaa taacatattt 1860
gaagagtctc atcggagaat tatgtccgtg aaaagaccag atgtcctaaa agctagactg 1920
tggattgagt ttgaatcaga gaaaggtctt gactatggg gtgtggccag agaattggtc 1980
ttcttactgt ccaaagagat gttcaacccc tactacggcc tctttgagta ctctgccacg 2040
gacaactaca ccttcagat caacccta atcaggcctct gtaatgagga tcatttgtcc 2100
tacttcaact ttattggaag agttgctggt ctggccgtat ttcattggga gctcttagat 2160
ggtttcttca ttagaccatt ttacaagatg atgttggga agcagataac cctgaatgac 2220
atggaatctg tggatagtga atattacaac tctttgaaat ggatcctgga gaatgacct 2280
actgagctgg acctcatgtt ctgcatagac gaagaaaact ttggacagac atatcaagt 2340
gatttgaagc ccaatgggtc agaaataatg gtcacaaaatg aaaacaaaag ggaatatatc 2400
gacttagtca tccagtggag atttgtgaac aggtccaga agcagatgaa cgccttcttg 2460
gagggattca cagaactact tcctattgat ttgattaaaa tttttgatga aaatgagctg 2520
gagttgtca tgtgcggcct cggatgatgt gatgtgaatg actggagaca gcattctatt 2580
tacaagaacg gctactgcc aaaccacccc gtcattcagt ggttctggaa ggctgtgcta 2640
ctcatggacg ccgaaaagcg tatccggtta ctgcagtttg tcacagggac atcgcgagta 2700
cctatgaatg gatttgccga actttatggt tccaatggtc ctgagctgtt tacaatagag 2760
caatggggca gtccctgagaa actccccaga gctcacacat gctttaatcg ccttgactta 2820
cctccatatg aaacctttga agatttacga gagaaacttc tcatggccgt ggaaaatgct 2880
caaggatttg aaggggtgga ttaagcacc tgtgcctcgg ggggtggtgt tcttcaagca 2940
agttctgct gcacttttg atttgcctaa cagacttttg cagaggcgat ggcagagagc 3000
agctgcaggc atggctccctg gagccgagcc ttcaccacgc aaaaaaaaaa aaaaaaaaaa 3060

```

<210> 9

<211> 623

<212> DNA

<213> Homo sapiens

<220>
 <221> modified_base
 <222> (460)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (524)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (542)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (555)..(556)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (571)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (594)
 <223> a, t, c, g, other or unknown

<220>
 <221> modified_base
 <222> (602)
 <223> a, t, c, g, other or unknown

<400> 9
 tgctgcaagt gacagggtcc aagaagcccc agggctcaga gctgaatgat gaagcgcagt 60
 ccccaaagtg cctggccacc cctccctccc tggatcactg ctgcctgggc ttgattgatt 120
 gattgattga ttgattgatt gatatttgaga gagattctca ctgtcaccca ggctggagta 180
 cagtgggtgcg atctcggtc actgcagcct ctgcctcccc gggtcaagca attctcctgc 240
 ctgagcctcc caagtagctg ggactacagg cagcgccac cacaccagc taattttgta 300
 tttttagtaa aagacggggt ttcaccatgt tgggccagga tggctctgat ctctgacct 360
 catgatccac ccgccccggc ttccaaagtg ctgggataca ggcataaacc cgacgcgccc 420
 agcatggaca ttttttttta atcccctgcc cttttcttgn ggcataattc attgcaggtc 480
 tcttctatac agatcatgga aaacacattt tcttaactga gttnttatta tttataccca 540
 gncacctcat gacannttta cctgtttaca nacaaaatgg gcacctgcca aaancaactt 600
 tnatataagg atgtccagg cct 623

<210> 10
 <211> 30
 <212> DNA
 <213> Unknown Organism

<220>
 <223> Description of Unknown Organism: Illustrative
 oligonucleotide

<400> 10
ctctgcaggc aaagccgtta ataaatattt

30

<210> 11
<211> 29
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<400> 11
tctgcaggca aagcccttaa taaaaattt

29

<210> 12
<211> 31
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<400> 12
caggtctagg acaacacagg ttaaaaagga c

31

<210> 13
<211> 31
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<220>
<221> modified_base
<222> (15)
<223> a, t, c, g, other or unknown

<400> 13
caggtctagg acaanacagg ttaaaaagga c

31

<210> 14
<211> 31
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<400> 14
acgggcagct cgttggaaaa tatagtaatg g

31

<210> 15
<211> 32
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<400> 15
gacgggcagc tcgtcggaaa atatagtaat gg

32

<210> 16
<211> 36
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<400> 16
acttgaaagg aatcacacca ttattcctta gaagta

36

<210> 17
<211> 33
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Illustrative
oligonucleotide

<220>
<221> modified_base
<222> (16)..(18)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (21)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (24)..(25)
<223> a, t, c, g, other or unknown

<220>
<221> modified_base
<222> (27)
<223> a, t, c, g, other or unknown

<220>

<221> modified_base

<222> (29)

<223> a, t, c, g, other or unknown

<400> 17

ttgaaaggaa tcaccnnnat nccnnanang taa

33